

Applicant therefore requests that the finality of that office action be withdrawn and further examination on the merits.

Rejections under 35 U.S.C. § 102

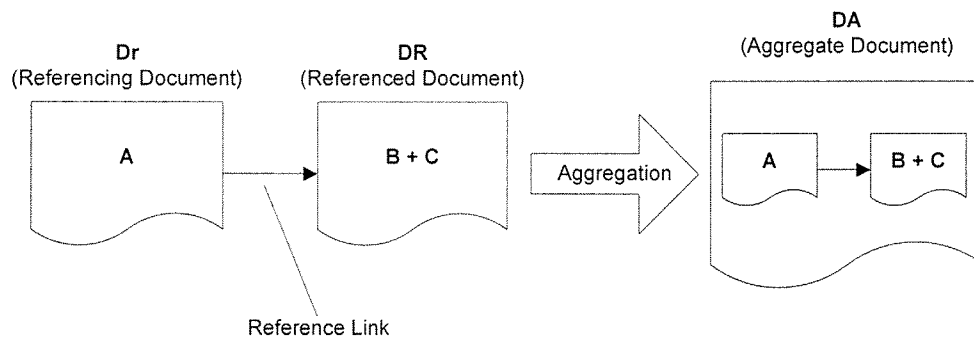
Claims 1, 10, and 14 stand rejected under 35 U.S.C. § 102(e) as being anticipated by US6643641 (Snyder). Applicant respectfully traverses the rejection.

Applicant claims a process for searching and a search engine for searching. This process allows searching among a collection of documents, the collection comprising a referencing document and a referenced document referenced in the referencing document (See claim 1).

For the sake of understanding only, the present invention is illustrated via an example.

Let **Dr** be a *referencing document* containing index term **A** and **DR** a *referenced document*, referenced in **Dr**, containing index terms **B** and **C**.

First, the process will aggregate a *referencing document* (**Dr**, containing index term **A**) and a *referenced document* (**DR**, containing index terms **B** and **C**) referenced in the referencing document to form an *aggregate document*. See claim 1. In the example illustrated below, this provides an aggregate document **DA**, which contains index terms **A**, **B** and **C**.



Then, the process will *index the aggregate document DA, based on index term contained in the referencing and referenced documents forming the aggregate document*. In other words, the process will index aggregate document **DA** for index terms **A**, **B** and **C**. This will

subsequently make it possible to retrieve aggregate document **DA** thanks to any and all of index terms **A**, **B** and **C**. In the absence of such aggregation in accordance with the present invention the referencing document **Dr** of the example would be retrieved only with the index term **A** and not with the index terms **B** and **C**, and the referenced document **DR** of the example would be retrieved only with the index terms **B** or **C** and not with the index term **A**, in contrast the aggregate document **DA** of the example would be retrieved with any of the index terms.

The index table, in the example considered here, would include:

Index term	Document
<b>A</b>	<b>DA</b>
<b>B</b>	<b>DA</b>
<b>C</b>	<b>DA</b>

After aggregation, the process will *search among aggregate documents* (See claim 1).

Finally, the process will *provide as a result an aggregate document* (See claim 1). For instance, let the search be a search for index terms **A** and **B** and **C**; this search will return, as a result, aggregate document **DA**. Neither the referencing document **Dr** nor the referenced document **DR** alone would satisfy that search criteria, nor would an index of the terms in each document which did not also index the terms in the aggregate document.

In contrast, prior art search engines do not disclose such features.

Patent US 6,643,641 to Snyder discloses a method which provides an abbreviated *representation* of searchable data files, in particular Internet/Intranet/Extranet html data pages, which represents *their text and linked graphics in a visual "snapshot form to supplement representations such as introductory text passages and URL addresses* (SUMMARY). Clearly, this aims at improving presentation of lists of results provided by search engines in response to a user query. That is, instead of solely providing *introductory text passages and URL addresses* of hits, here Snyder allows for aggregating graphics to the results.

However, all this concerns how search results are rendered to users and does not particularly relate to the indexing step (which is upstream the user search query and search report). This is described in Snyder, Col. 14, second and third paragraphs. In particular, the third paragraph describes a mechanism in which graphics are displayed with search results in the search report (see FIG. 3, ref. 80, 86, 82, 35). Creation of the snapshot does not modify the result of the search. Therefore, the creation of the snapshot in Snyder has poor resemblance with the aggregation recited in the claims..

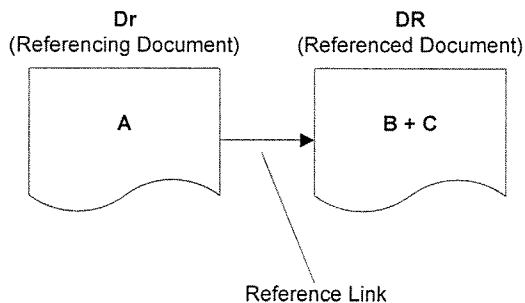
More specifically, the Examiner has cited the following passages.

1. Col. 2, line 59 to col. 3 line 8: relates to prior art. In particular it defines conventional index database of the known art, which may include various features.

One may claim that such databases looks like aggregated documents. However, such databases are the result of a classical known indexing step, which are of no particular interest with respect to the present invention. In particular, such known database or indexing fails to disclose a step of “indexing an aggregate document, based on index terms contained in the referencing and referenced documents forming the aggregate document”.

Indeed, a usual database is built of a set of index tables. However, the index tables contain index terms, and for each index terms, a list of documents which contain this index term.

In the example considered above, the index table of a usual database would be:



Index term	Document
<b>A</b>	<b>Dr</b>
<b>B</b>	<b>DR</b>
<b>C</b>	<b>DR</b>

Accordingly, one sees in the above table that index terms **A**, **B** or **C** relate to physical document Dr or DR. they never relate to either a physical or even a logical document that would be an aggregated document. This markedly differs from the result typically obtained with present invention as to the index table, which would include:

Index term	Document
<b>A</b>	<b>DA</b>
<b>B</b>	<b>DA</b>
<b>C</b>	<b>DA</b>

Thus, even considering the implicit disclosure of every database system, this does not disclose the claimed step of *indexing*. Instead, databases of the prior art are the result of indexing physical documents and not aggregated documents.

2. Col. 9, lines 57–61 describes typical data included in web pages, including addressed files. However, this does not disclose *aggregating a referencing document and a referenced document* and then *indexing an aggregate document* (based on index terms contained in **both** documents forming it).

On the contrary, next paragraph recites, “The collected information from downloaded files, particularly text files, is processed according to a **generally conventional text processing or categorizing technique** 68 to build a text or descriptor index in database 62” (col. 9, line 62, emphasis added).

3. With respect to a search report, col. 14 lines 22–40 describes “encoding and providing in the search report . . . a standardized graphic representation . . . of the appearance and rendering of each page at the time that the page is indexed.” It refers to this standardized graphic representation as a snapshot.

In particular: “The snapshot is acquired when the page is initially loaded by the crawler 60 for indexing (FIG. 2). The snapshot is rendered, converted to the compressed format and

stored.” The paragraph goes on with “When the subject page is selected in a search (FIG. 3), transmitted to the user are the individual snapshots, which have been stored locally to the search portal processor 78, in association with the index/categorization database. In this way the snapshots 35 of the hit page (which may be one of a number of hits that are reported to user 30) **is shown when providing the search report.**” (emphasis added.) Note, however, it is the individual component pages that are indexed and searched, not the snapshot as an aggregate. Moreover, the snapshot as described in Snyder appears to be a graphical representation which does not include the text and metadata contained in the source documents and thus cannot be indexed as an aggregate document in accordance with the subject matter claimed by the applicant herein. There is no aggregation in the sense of the invention and no indexing of an aggregate document.

4. Col. 16 lines 52–58: describes rendering a page layout including graphics, with little interest concerning the present invention.

5. Col. 24 lines 9–20, discusses how the crawling operates. Here a specific technique is involved in order to prevent the crawler (web agent) from moving on a next web site before it has completed a current web site (the problem is explained in col. 23, lines 55–60). That technique, however, concerns what is done during the crawling, not during the indexing.

To clarify, crawling and indexing are two different steps in search engine techniques. In Snyder, after crawling the pages (agent A, 66, top left part of FIG. 2), web page ingredients are stored in a buffer 92 (see FIG. 2) for subsequent indexation 68. However, that indexation relies on conventional techniques, as mentioned above.

Accordingly, care should be taken not to confuse teachings of Snyder which relate to different moments/steps like crawling, indexing, searching, rendering results to user, etc. In this

respect, while items 3 and 4 have something in common, they do not concern steps occurring at the same time as item 1 or 5. So, it does not make sense to combine the disclosures of said items when interpreting the disclosure of Snyder.

Therefore, the passages cited by the Examiner are not relevant to the present invention.

Technically, Snyder describes:

- crawling the web;
- storing files in a buffer 92;
- capturing snapshots 35 stored in the buffer (see agent B, ref. 95 in FIG. 2)
- independently from capturing snapshots, conventionally indexing 68 files stored in the buffer (see FIG. 2, wherein capturing snapshots and indexing occur in different routes) ;
- searching among indexed files; and
- providing, as a result of the searching step, an aggregate document which comprises snapshots aggregated in the search report.

Accordingly, it does not disclose the claimed steps of:

- aggregating a referencing document and a referenced document referenced in the referencing document to form an aggregate document;
- indexing the aggregate document based on index term contained in the referencing and referenced documents forming the aggregate document.

A summary comparison of the claimed subject matter to the particular disclosure of Snyder further confirms that Snyder does not teach or suggest the subject matter as claimed.

With regard to Claim 1, Snyder does not teach or suggest aggregation according to the invention (that is, before indexing). In particular, Col. 9, lines 48–61, describes using multiples processes to retrieve files into a buffer “so that one file can be processed while waiting to

retrieve another”, but significantly does not teach or disclose aggregating documents and indexing said aggregate documents. Col. 16, lines 52–58, describe rendering an html page but contain no disclosure as to aggregation or indexing. Col. 24, lines 9–20, disclose preferential queuing of different files which are not aggregated and are indexed separately, but does not teach or suggest aggregating the documents and indexing the aggregate document.

Further with regarding to Claim 1, Snyder does not teach or suggest indexing the aggregate document. Again, Col. 24, lines 9–20, disclose preferential queuing of different files which are never aggregated and are indexed separately, but does not teach or suggest aggregating the documents and indexing the aggregate document. Similarly, col. 14, lines 22–40, disclose creation of a graphical “snapshot” which is in the nature of a thumbnail depiction, but the snapshot is not an aggregate document in the sense of the invention because it does not include terms from the source documents to be indexed and, moreover, the “snapshot” of Snyder is not separately indexed apart from the indexing of the referenced and the referencing document.

Further with regard to Claim 1, Snyder does not teach or suggest searching by operating on the aggregate document index because it does not create an index associating index terms with aggregate documents. Snyder describes, in col. 2 line 59 through col. 3 line 8, searching an index separately including terms associated with referencing documents and, if also indexed, terms associated with referenced documents, but does not teach or suggest an aggregate document index associating terms from both with the same aggregate document.

Further with regard to Claim 1, Snyder does not teach or suggest “returning . . . a located aggregate document.” Snyder describes in col. 10 lines 17–38 a search that “reports . . . web pages that met the search criteria **in a conventional manner.**” The report includes a graphical

representation of web page fully rendered, but the search itself does not return an aggregate document in accordance with the subject matter claimed herein.

Claims 10 and 14 of the application depend from Claim 1, and are distinguished from Snyder based on the arguments set forth above.

For all the reasons above, the claimed process is novel over Snyder. Withdrawal of the rejection is respectfully requested.

Rejections under 35 U.S.C. § 103

Claims 8–9 stand rejected under 35 U.S.C. § 103(a) as being obvious over U.S. Pat. No. 6,643,641 to Snyder in view of US 2002/0052894 by Bourdoncle. Claims 11–13 stand rejected under 35 U.S.C. § 103(a) as being obvious over U.S. Pat. No. 6,643,641 to Snyder in view of the publication *Google Hacks*. Applicant respectfully traverses these rejections.

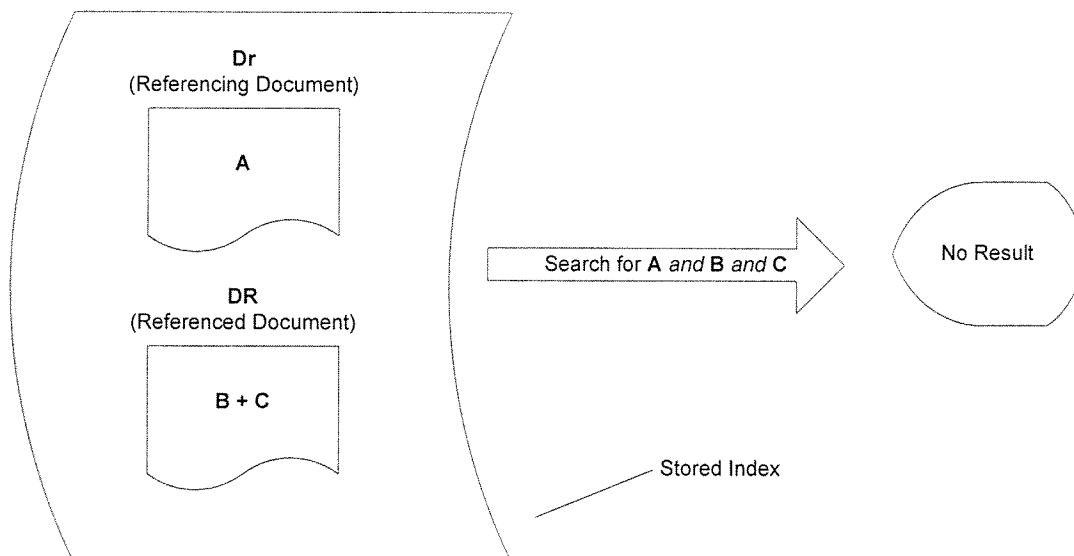
The claimed subject-matter is non-obvious over Snyder, in view of either Bourdoncle or *Google Hacks*.

Snyder describes the capture and rendering of composite snapshots. However, Snyder provides no detail as to the indexing. It is believed that Snyder employs a conventional indexing and search methodology, as suggested in e.g. in col. 9, line 62, as discussed above.

One difference between the claimed invention and the disclosure of Snyder is that the prior art system of Snyder will not return the same hits as the claimed invention.

In the example given above, if a user requests a search containing information **A and B and C**, Snyder's search engine will give no solution, because none of the referencing or referenced documents individually contain **A and B and C**. In particular, neither the referencing document **Dr** nor the referenced document **DR** contains all of the relevant search terms.

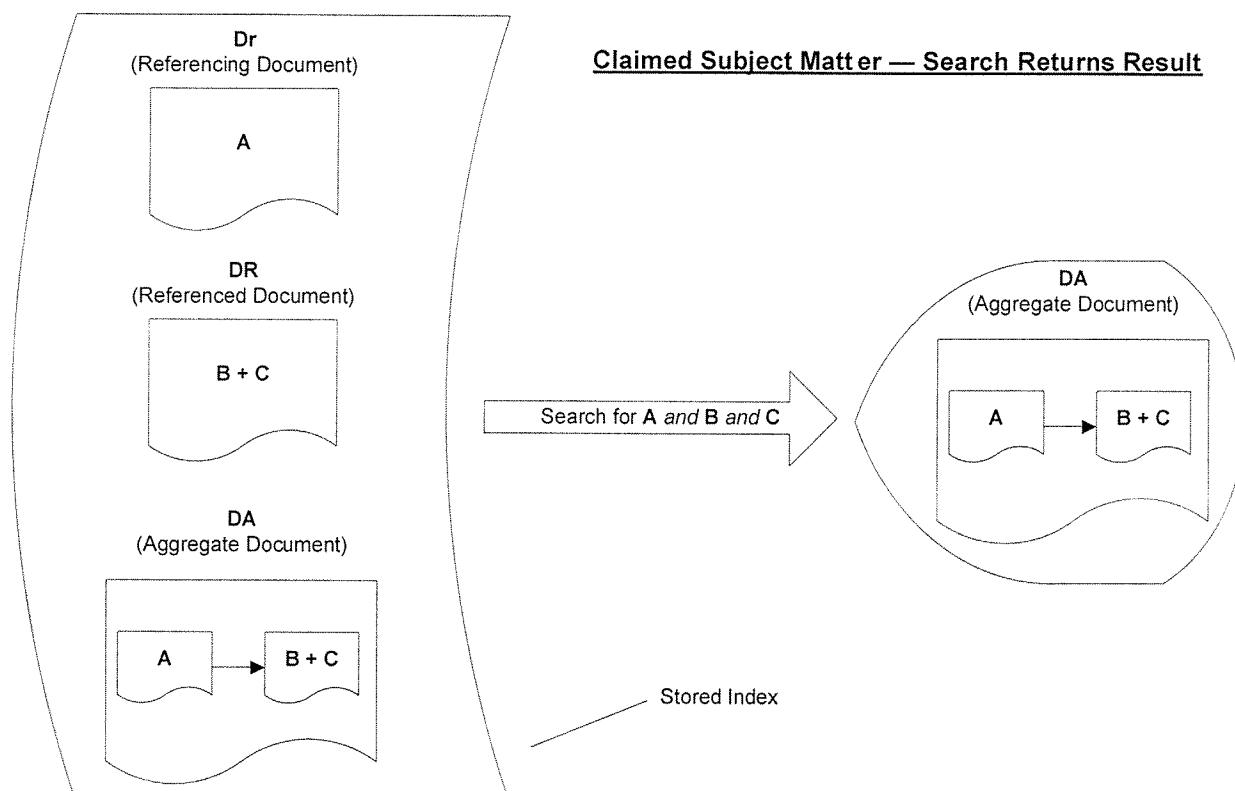




By contrast, in the subject matter now claimed by applicant, the index of the database of the search engine, which indexes aggregate documents, would include:

Index term	Document
<b>A</b>	<b>DA</b>
<b>B</b>	<b>DA</b>
<b>C</b>	<b>DA</b>

Therefore, if a user requests a search in accordance with an embodiment of the subject matter claimed herein containing the information: **A and B and C**, the search engine returns one hit, the aggregate document **DA**, because there is a step of indexing aggregate documents. The documents were aggregated before indexing, and the indexing step itself occurred before search queries.



Therefore, the invention differs from Snyder at least in that it provides the step of aggregating a referencing document and a referenced document referenced in the referenced document to form an aggregate document and the step of indexing the aggregate document, based on index terms contained in the referencing and referenced documents forming the aggregate document.

In particular, the claimed solution makes it possible to retrieve results, thanks to keywords contained in documents that are different but linked, as exemplified above. Snyder does not provide a solution to this problem and therefore would not be considered by a person of ordinary skill in the art as a solution to this problem.

Furthermore, Snyder teaches away from the invention because:

- Snyder discloses relying on a conventional indexing technique, in contrast with the present invention;

- Snyder suggests forming a composite of textual and non-textual documents in a non-textual graphical form, which does not solve the technical problem of the invention.

More particular with respect to Claim 8 and Claim 9, Snyder in view of Bourduncle does not teach or suggest the subject matter of either Claim 8 as a whole or Claim 9 as a whole. The examiner states that Snyder does not disclose the use of an inverse index. More significantly, however, Snyder does not disclose an indexable aggregate document. The “snapshot” of Snyder is a graphical thumbnail representation of a rendered web page, and teaches away from an indexable aggregate document in accordance with the subject matter claimed by applicants. There is no motivation to combine the reverse index of Bourduncle with the snapshot of Snyder, nor is there an reasonable expectation that one could successfully do so.

More particularly with respect to Claims 11, 12, and 13, Snyder in view of *Google Hacks* does not teach or suggest the subject matter as a whole as recited in the claims. In particular, the ability to search for referenced documents in the form of images as described in *Google Hacks* does not teach or suggest a search of aggregate documents including a subset of referenced documents.

Therefore, Applicant believes that the present invention is both novel and non-obvious over Snyder and the other cited references. For these reasons, applicant respectfully requests withdrawal of the rejection for obviousness.

#### Closing Remarks

For the foregoing reasons, applicant submits that the proposed amendments would place the subject application in condition for allowance and earnestly solicits an early Notice of Allowance. **Should the Examiner be of the opinion that a telephone conference would**

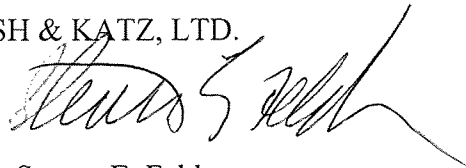
**facilitate prosecution of the subject application, the Examiner is respectfully requested to call the undersigned at the below-listed number.**

A Petition for an Extension of Time of one month and its required fee are enclosed to permit the Examiner extra time to consider this paper. The Commissioner is hereby authorized to charge any additional fee which may be required for this application under 37 C.F.R. §§ 1.16-1.18, including but not limited to the issue fee, or credit any overpayment, to Deposit Account No. 23-0920. Should no proper amount be enclosed herewith, as by a check being in the wrong amount, unsigned, post-dated, otherwise improper or informal, or even entirely missing, the Commissioner is authorized to charge the unpaid amount to Deposit Account No. 23-0920. Further, should any additional petition be necessary, Applicant requests that the currently filed paper constitute any such necessary petition.

Respectfully submitted,

WELSH & KATZ, LTD.

By



Steven E. Feldman  
Registration No. 40,604

July 17, 2007  
WELSH & KATZ, LTD.  
120 South Riverside Plaza  
22nd Floor  
Chicago, Illinois 60606  
(312) 655-1500